Final report of the NKFI PD 121322 grant (2016-2019) titled:

'Investigating the transcriptional regulation of high molecular weight prolamins of cereals'

Csaba Éva

Agricultural Institute, Centre for Agricultural Research

Martonvásár, Hungary

1) Yield increase: improving photosynthetic efficiency

Environmental stress can reduce the potential yield of crop plants by as much as 70% (Agarwal et al. 2006), presenting a major challenge in our quest for sustainable food production. The issue could not be more relevant in the age of climate change, which explains a third of global crop yield variation (Ray et al. 2015). However there is a huge pressure on breeders and biotechnologists these days to maintain or even increase crop yields at least to compensate for the increase of world human population, losses of agricultural land due to urbanization, soil degradation, and the growing demand for food crops as energy-sources (reviewed by Éva et al. 2019). During this project various approaches were tested and suggested to improve the quality and yield of cereals. To meet these goals we had initially planned a comprehensive transcriptome sequence analysis of heat- and drought-stressed barley plants. Our intention was to identify transcription factors and other regulators which may provide protection against heat and drought stress and help to maintain the yield stability under such adverse conditions. Since our NKFI-6 No. 119891 grant application was not approved, the cost of those expensive experiments could not be covered, and a different approach has been pursued instead. A review paper was published on photosynthetic improvement and proposed as a way of increasing cereal yield (Éva et al. 2019). The review contains massive amounts of literature data to support recommendations on the improvement of photosynthetic efficiency. Other main elements of plant productivity, such as harvest index, life cycle/senescence and stress tolerance, were also briefly discussed. We emphasised the connection between these traits and suggested two approaches to improve several factors simultaneously. First, the introduction of sugar alcohol metabolism to cereals (utilising sorbitol or mannitol as the main phloem-transported photosynthetic product instead of sucrose) was proposed, which could lead to higher photosynthetic activity, efficient source to sink transport and improved osmotic stress tolerance. This metabolic route was already described for several natural plants including celery, ash tree, and members of the Rosaceae family (Pharr et al. 1995). We also called for an ambitious plan to incorporate some form of carbon dioxide transport from the root to the shoot system within the xylem (**Figure 1**). It would also include sugar alcohol metabolism as a means for improved energy supply of the root system. We believe that while water deficit is one of the most prevalent global stressor limiting plant productivity, our scheme could promise a high photosynthetic activity with

extreme drought tolerance. This route is unexplored yet in the scientific literature as well as in nature. Further research should clarify its feasibility or natural occurrence.

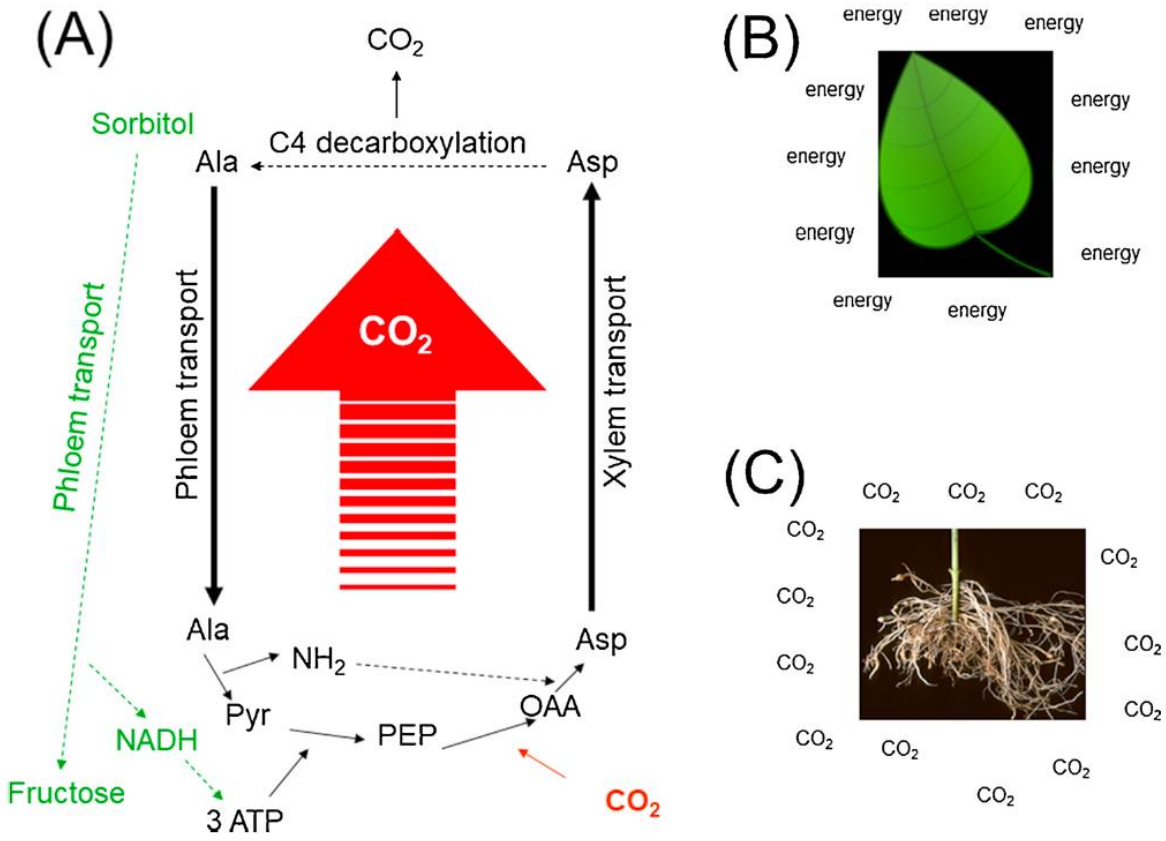


Figure 1. A proposed scheme of root to shoot carbon dioxide (CO_2) transport, through NAD ME or PEPCK-type C4 photosynthesis split between the two organs. Our scheme (A) is intended to transfer energy (through sugar alcohol metabolism) from the energy-rich leaves (B) to the energy-poor roots (C) in exchange for CO_2, which is abundant in the vicinity of roots. CO_2 is fixed by PEP-carboxylase in roots, yielding oxaloacetate (OAA). Oxaloacetate is transaminated to form the more stable aspartic acid (Asp), which is then transported to the shoot system through the xylem. Aspartic acid is converted to alanine (Ala) and CO_2 in leaves through either NAD-ME or PEPCK C4 photosynthesis process. The produced CO_2 is fixed by Rubisco. Alanine is transported back to the roots through the phloem. Pyruvate (Pyr) is formed from deamination of alanine. Phosphoenolpyruvate (PEP) is regenerated from pyruvate at the expense of two ATP molecules. The approach may be supported by the implementation of sugar alcohol metabolism. It would involve sorbitol supply of roots (green) which is more energy rich per carbon atom, compared to the common sucrose. Sorbitol is degraded to fructose by sorbitol-dehydrogenase, forming fructose and NADH. Reducing power of a NADH molecule is eligible to produce three ATP molecules. (Source: Éva et al. 2019)

2A) Epigenetic regulation of wheat HMW GS genes

In the other parts of the project we focused on molecular aspects of the nutritional quality of bread wheat. Wheat breeders have long focused on improving bread-making quality. The dual aims of increasing yield and improving quality are interconnected because once high quality wheat has been grown, no further land has to be used for growing improved wheat. High molecular weight glutenin subunit (HMW GS) proteins are major determinants of baking quality, because they provide the wheat dough with elasticity (Shewry and Halford 2002; Delcour et al. 2012). The regulation of HMW GS gene expression has been summarised in earlier works of our team (Makai et al. 2014, 2015). HMW GS protein accumulation is confined to the starchy endosperm, in accordance with the protein's function of supplying nutrients to the germinating plantlet. The expression of the encoding genes is induced primarily at 5 days after anthesis, and the role of ABA hormone should be underlined here. The expression is also regulated by environmental factors like water, nitrogen and sulphur availability. These signals affect HMW GS gene expression primarily through transcriptional regulation. According to a recent result at the beginning of this project (Guo et al. 2015), however, the expression of HMW GS genes during endosperm development is activated by histone acetylation, a major type of epigenetic regulation. Since the regulation of HMW GS genes was previously thought not to be regulated by epigenetic factors, we designed experiments to validate that report. These works was performed in collaboration with László Tamás's laboratory (ELTE University, Budapest), where transgenic barley plants harbouring HMW GS promoter::GUS (β-glucuronidase, *uidA*) reporter gene constructs (p*Glu-1Bx7*::*uidA*::*nos*T and p*Glu-1By9*::*uidA*::*nos*T) had already been produced. During germination, these plants were treated with Trichostatin-A (TSA), a potent histone deacetylase inhibitor causing hyperacetylation of histone proteins, which activates genes regulated by histone acetylation (Tanaka et al. 2008). It was found that *uidA* expression, confirmed by qPCR and GUS-staining (**Figure 2**), was greatly increased in vascular tissues of the coleoptile as a result of TSA-treatment. This observation confirmed the findings of Guo et al (2015) in the sense that HMW GS genes are inhibited in non-endosperm tissue by the lack of histone acetylation. The novel result was that this type of regulation might be more characteristic for the y-type than for the x-type HMW GS genes. We also attempted to purify transcription factors (TF) binding HMW GS gene promoters by biolistic transformation of biotinylated HMW GS gene promoter fragments to wheat leaves and subsequent re-extraction of protein-crosslinked probe (DNA affinity purification), but were only able to purify HMW GS gene promoter-binding histones. Mass spectrometry analysis indicated, that the N-terminal lysines of H3 histones were not acetylated in the leaf. It confirmed our results about epigenetic inhibition HMW GS gene expression outside the target organ. Our findings on epigenetic regulation of HMW GS genes have already been published (Éva et al. 2018).

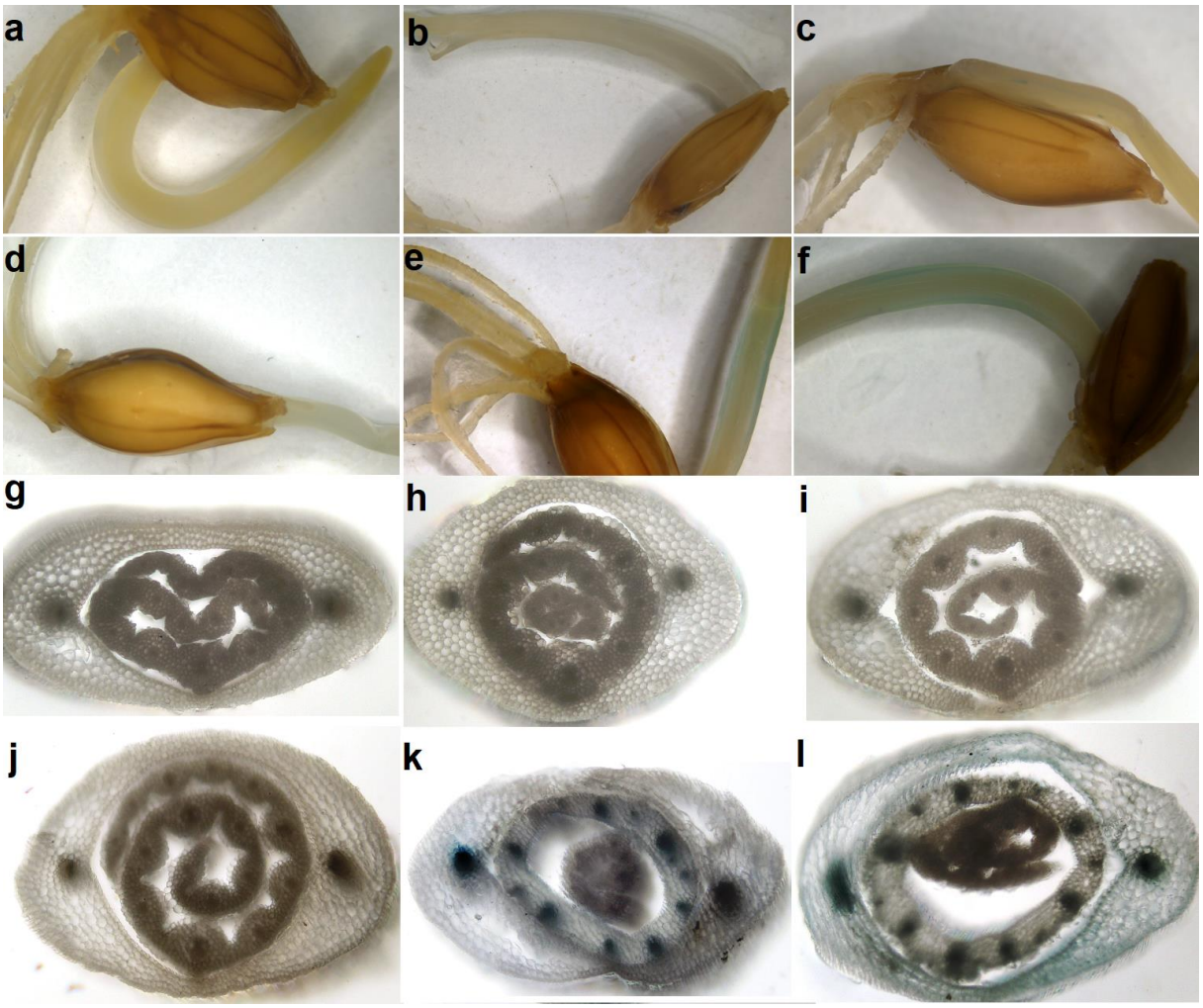


Figure 2. Effects of TSA treatments on GUS expression in non-transgenic and HMW GS gene (*Glu-1Bx7* and *Glu-1By9*) promoter::*uidA*-reporter gene containing transgenic 'Golden Promise' barley seedlings and coleoptiles. Seedlings untreated: non-transgenic control (a), p*Glu-1Bx7*::*uidA* (line II) (b) and p*Glu-1By9*::*uidA* (line II) (c); TSA-treated: non-transgenic control (d), p*Glu-1Bx7*::*uidA* (e) and p*Glu-1By9*::*uidA* (f). Bisected coleoptiles untreated: non-transgenic control (g), p*Glu-1Bx7*::*uidA* (h) and p*Glu-1By9*::*uidA* (i); TSA-treated: non-transgenic control (j), p*Glu-1Bx7*::*uidA* (k) and p*Glu-1By9*::*uidA* (l).

2B) Transcriptional regulation of wheat HMW GS genes

While epigenetic factors clearly influence the expression of HMW GS genes (as shown above), transcriptional regulation is considered the most important way of their regulation. It is mostly a result of binding transcription factors to HMW GS gene promoter elements. We identified a (second) *cis*-regulatory module (CRM2) in HMW GS gene promoters, which acts as an inhibitory element and exerts tissue-specific control on HMW GS gene expression (**Figure 3**). A promoter mutant was constructed with mutant CRM2 (Myb and VP1 TF-

binding sites mutated, only a bZIP TF binding site remaining intact). Its activity in leaf followed that of the wild type indicating that perhaps a bZIP was the inhibitory TF that bound this element in leaves. Other promoter-mutants have also been studied. The basal (177 bp length) promoter region of the *Glu-1Bx7* HMW-GS gene promoter was still active and showed higher activity in leaves compared to the wild type probably because of the lack of inhibitory sequences present in longer promoters (**Figure 3**). Interestingly, a double mutant version without either CRM2 or ABRE|CBF (another inhibitory module) showed no activity in leaves. This maybe a result of altered promoter structure which is important during DNA-bending hypothesized for HMW GS gene promoters (Makai et al. 2015). It also has to be noted that while these endosperm-specific promoters leaked somewhat in leaves, their activity was much higher in their native endosperm. In that tissue, the wild type promoter was the strongest (data not shown).

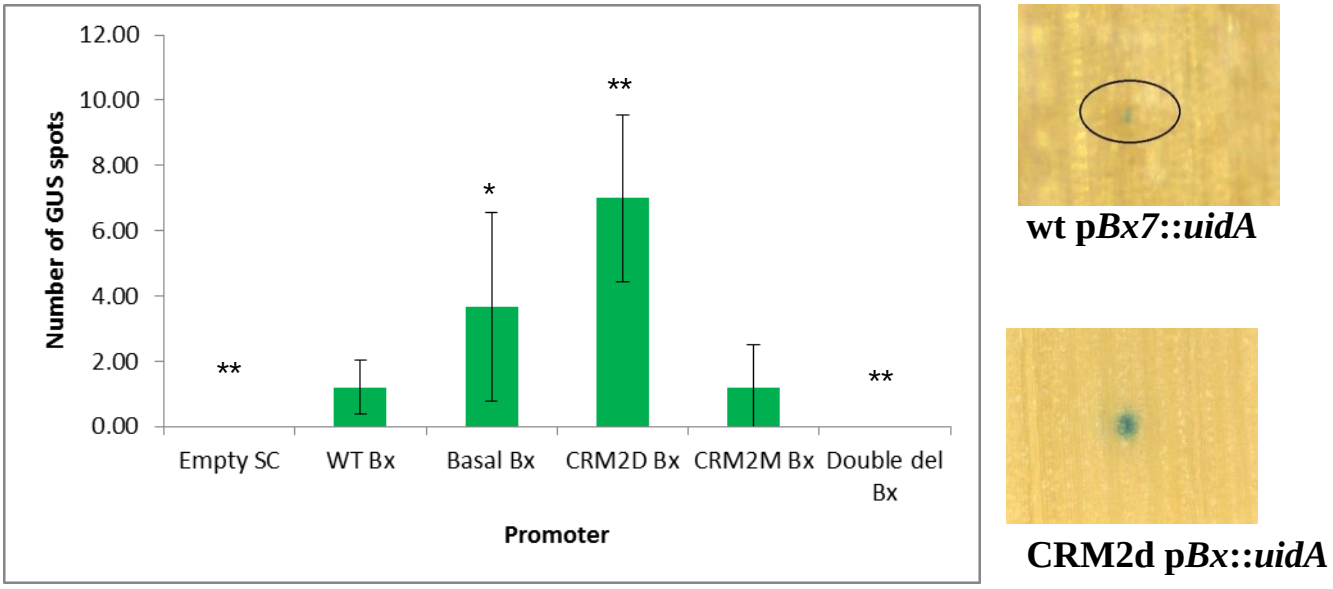


Figure 3. Activity of the empty vector (termed ShortCAMBIA, SC, our own derivative of pCAMBIA1391z with the bacterial *lac* promoter removed), wtBx7, basal region of *Glu-1Bx7* gene promoter (basal) and *Glu-1Bx7* gene promoter-mutants (CRM2del, CRM2M) in wheat leaves. No. of replicates = 5. Significant differences from wtBx7 are indicated by * (95% level) or ** (99% level).

We also took steps to identify the CRM2-binding bZIP TF by biolistic transformation of biotinylated HMW GS gene promoter fragments to wheat leaves and subsequent re-extraction of protein-crosslinked probe (DNA affinity purification). These efforts only lead to the isolation of histones as previously mentioned, therefore CRM2-region binding TFs were studied indirectly. To this end, wheat bZIP TF coding sequences were cloned from 'Bánkúti 1201' wheat leaf cDNA. Based on the ExpVip database (http://www.wheat-expression.com/), one member showing high expression in leaves was chosen for eight of the 10 bZIP

phylogenetic groups known in wheat, where leaf-specific members exist (Li et al. 2015). The cloned genes were later used in trans-activation experiments (transient expression of HMW GS gene promoter driven *uidA* gene) as antisense constructs (see **Figure 4a**).

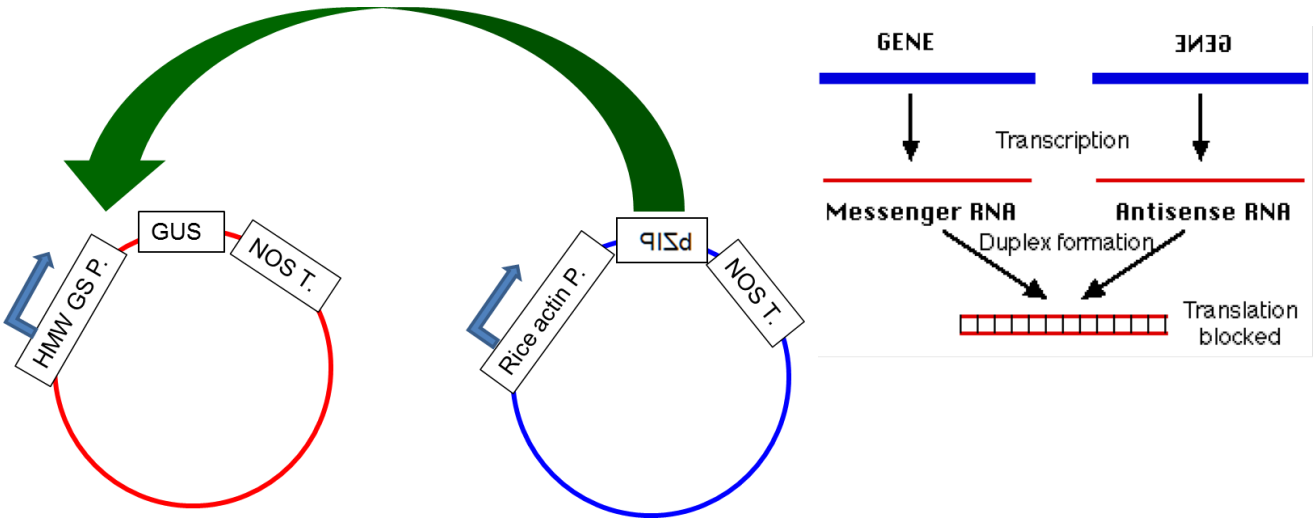


Figure 4a

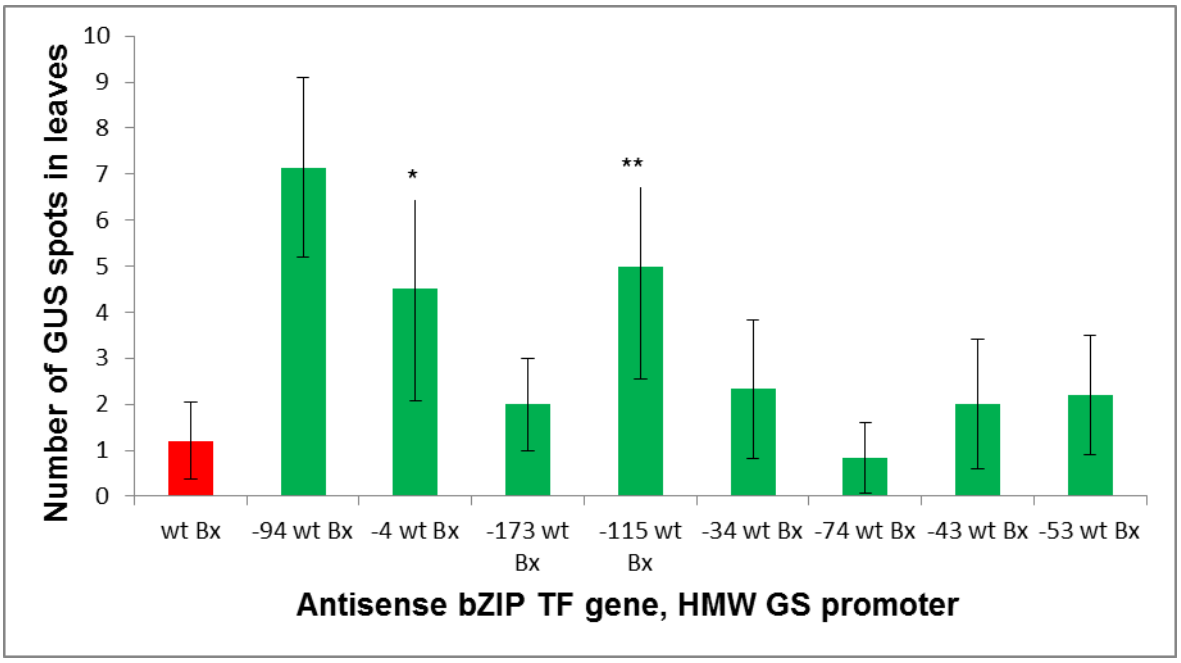


Figure 4b

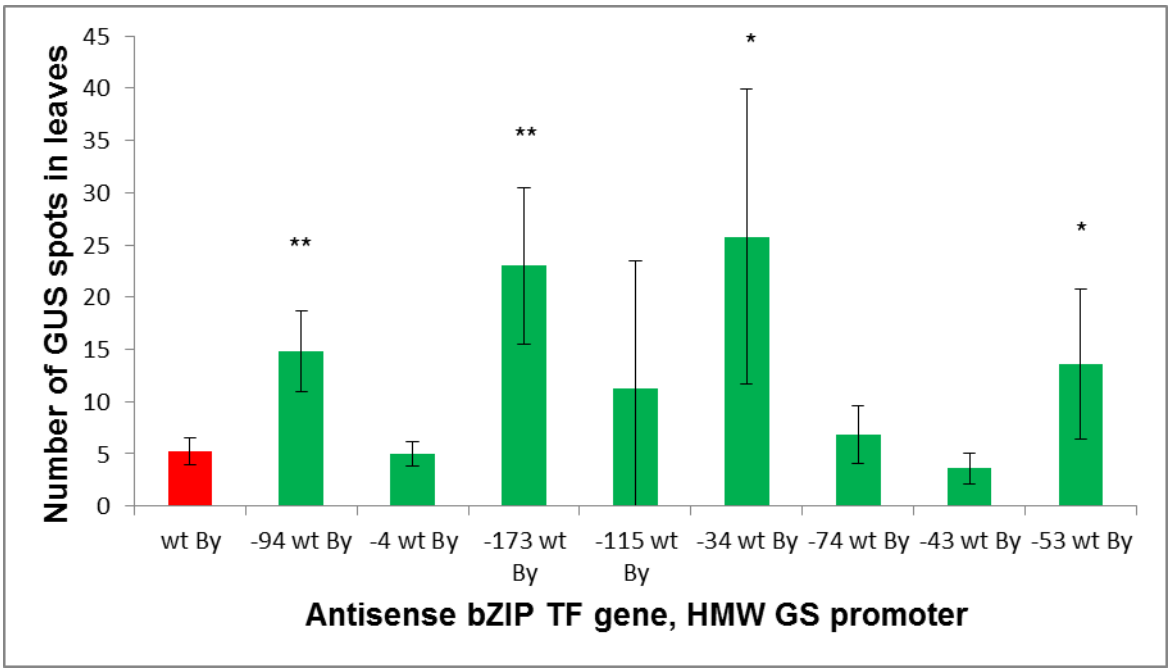
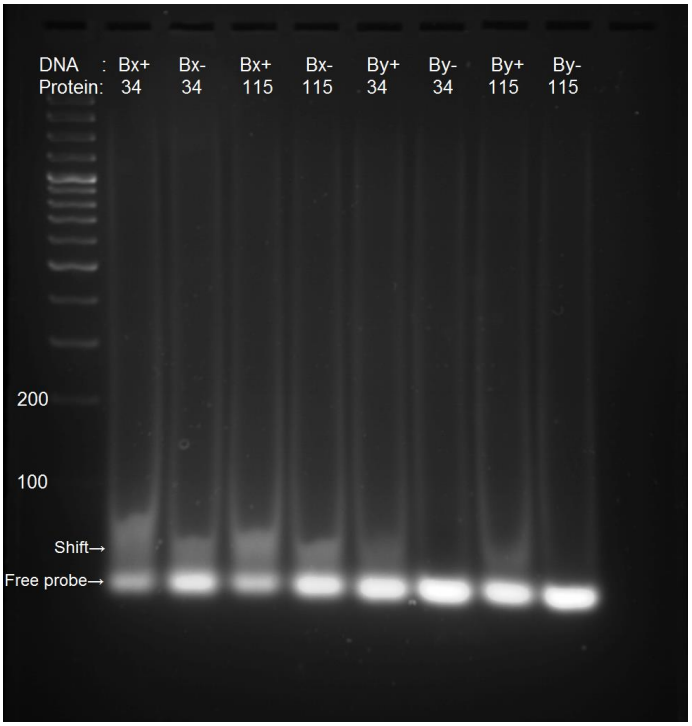


Figure 4c

Figure 4. Scheme of trans-activation assays for studying CRM2-binding transcription factors (a) and results of such trans-activation assays on CRM2 region of *Glu-1Bx7* (b) and *Glu-1By8* (c) HMW GS gene promoters in 'Bánkúti 1201' wheat leaves. Leaf-specific bZIP transcription factors (based on ExpVIP database) were used as antisense constructs. Significant differences from the wild-type gene promoter alone are indicated by * (95% level) or ** (99% level).

As seen on **Figure 4b and c**, the trans-activation experiments yielded many TF candidates, which could repress *Glu-1Bx* (TabZIP94, TabZIP4, TabZIP115) and *Glu-1By* (TabZIP94, TabZIP173, TabZIP34, TabZIP53) type HMW GS genes in wheat leaves. Among these, we consider TabZIP34 and TabZIP115 as the strongest hits. These factors are also part of the early genetic programme, expressing at early phase (0-5 days post anthesis) of endosperm development (Éva et al. 2018). HMW GS genes do not express at this stage, therefore the early programme may contain suppressor factors, so the regulation may be similar to that of vegetative tissues. This programme was identified by our co-expressional network bioinformatic analysis aiming to identify genetic programmes and key TF genes that regulate HMW GS genes (Éva et al. 2018). Binding of TabZIP34 and TabZIP115 to the CRM2 region of HMW GS gene promoters was also validated using electrophoretic mobility shift assay (EMSA), based on a published method (Holden and Tacon 2011). His-tag fusion TabZIP34 and TabZIP115 were ectopically expressed in and purified from *E. coli*. These two proteins reacted with 25 bp dsDNA probes from the CRM2 region of *Glu-1Bx7* and *Glu-1By8* gene promoters, respectively. Both wild-type sequence probes (Bx+, By+) containing the predicted bZIP binding site TGTCAT, and mutated probes (Bx-, By-) containing the TGTTTT

sequence, were tested. All reactions contained a poly(dI-dC) competitor as well. The results can be seen on **Figure 5**: both TFs could bind to both promoters. In case of the *Glu-1By* gene promoter, they bound only to the wild type probe. In case of the p*Glu-1Bx* gene promoter, they also bound to the mutated probe, though with less affinity.



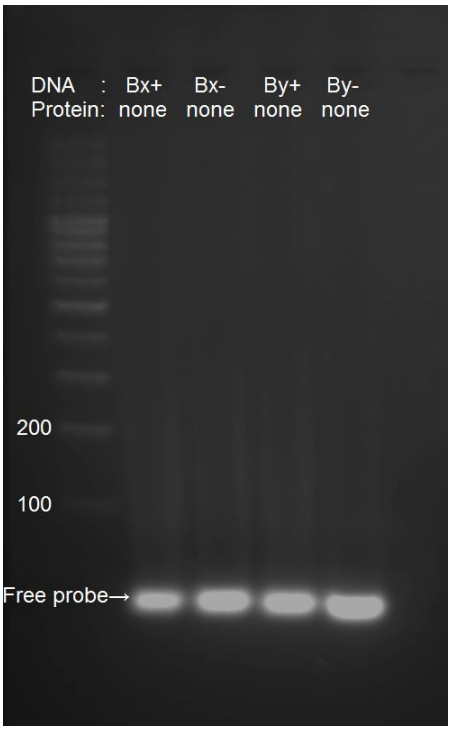


Figure 5a **Figure 5b**

Figure 5. EMSA validation of the binding of TabZIP34 and TabZIP115 TF proteins to 25 bp dsDNA probes from the CRM2 region of HMW GS gene promoters (*Glu-1Bx*, *Glu-1By*). Bx+ and By+ probes included the predicted bZIP binding site TGTCAT, while the mutant probes (Bx-, By-) contained the sequence TGTTTT. All reactions contained a poly(dI-dC) competitor. Protein containing reactions can be seen on **Figure 5a**, while protein-free reactions on **Figure 5b**. Shifts indicate protein binding.

The results regarding the inhibitory nature of CRM2 and bound factors will soon be published (Éva et al. manuscript) and will replenish existing literature data. Inhibiting effect of bZIP transcription factors is rather remarkable. Earlier works showed that bZIPs like the storage protein activator (SPA) bZIP activate HMW GS gene expression in the endosperm (Ravel et al. 2009), while Myb factors repress storage protein expression in vegetative tissues (Suzuki et al. 1998; Diaz et al. 2002; Chen et al. 2005). However, the regulation of HMW GS genes may be different from other storage proteins. The repression by Myb factors has not been reported on HMW GS gene promoters to our knowledge. When we mutated a Myb binding site in CRM2, it had no effect on the HMW GS gene promoter strength (see **Figure 3**). Our present results show that inhibitory bZIPs cause the repression in vegetative tissues, binding

to the TGTCAT motif. This motif was already suggested to be involved in auxin-induced downregulation (Zemlyanskaya et al. 2016). Recently, TabZIP34 (called SPA Heterodimerizing Protein) was also identified as a repressor of HMW GS gene expression (Boudet et al. 2019), though the authors only examined the endosperm. It has to be noted, that according to the ExpVip database, both TabZIP34 and TabZIP115 show much higher expression level in leaves compared to the starchy endosperm, therefore their role as repressor may be more pronounced in vegetative tissues. Based on the literature and our own data, we hypothesize that the expression of HMW GS genes is repressed in vegetative tissues and early stage endosperm (0-5 days post anthesis) by auxin-induced downregulation, mediated by TabZIP34 and TabZIP115 alongside epigenetic factors like histone deacetylation. HMW GS genes may be induced during late endosperm development by ABA signal through the action of SPA bZIP and Gamyb transcription factors (Ravel et al. 2009; Guo et al. 2015). This latter has been shown to recruit histone acetyltransferases for epigenetic activation of HMW GS genes (Guo et al. 2015). The expression may also be mediated at later stages by TabZIP34 and TabZIP115. While these mechanisms may at least partially account for the regulation of HMW GS, many unresolved questions remain. For example, HMW GS gene promoters not only contain binding sites for bZIP and MYB, but also for DOF, B3/VP1, CBF and NAC TFs. Their effect on HMW GS gene expression would be worth to study as these TFs were shown to modulate the expression of other cereal seed storage proteins (Moreno-Risueno et al. 2008; Gupta et al. 2011; Mathew et al. 2016).

The long-term benefit and consequence of this research may be the breeding for transcription factor genes with positive effect on HMW GS gene expression and breeding against inhibitory TFs. It could lead to increased HWM GS gene expression in the endosperm and thereby result in improved quality of bread wheat.

28 October 2019, Martonvásár

Csaba Éva

project leader

3) References

Agarwal PK, Agarwal P, Reddy MK, Sopory SK (2006). Plant Cell Rep 25:1263–1274. https://doi.org/10.1007/s00299-006-0204-8

Boudet J, Merlino M, Plessis A, et al. (2019) The bZIP transcription factor SPA Heterodimerizing Protein represses glutenin synthesis in *Triticum aestivum*. Plant J 97:858–871. https://doi.org/10.1111/tpj.14163

Chen R, Ni Z, Nie X, et al. (2005) Isolation and characterization of genes encoding Myb transcription factor in wheat (*Triticum aestivum* L.). Plant Sci 169:1146–1154. https://doi.org/10.1016/j.plantsci.2005.07.018

Delcour JA, Joye IJ, Pareyt B, et al. (2012) Wheat gluten functionality as a quality determinant in cereal-based food products. Annu Rev Food Sci Technol 3:469–492. https://doi.org/10.1146/annurev-food-022811-101303

Diaz I, Vicente-Carbajosa J, Abraham Z, et al. (2002) The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. Plant J 29:453–464. https://doi.org/10.1046/j.0960-7412.2001.01230.x

Éva C, Oszvald M, Tamás L (2019) Current and possible approaches for improving photosynthetic efficiency. Plant Sci 280:433–440. https://doi.org/10.1016/j.plantsci.2018.11.010

Éva C, Szőke-Pázsi K, Makai S, et al. (2018) In vivo DNA affinity purification and histone deacetylase inhibitor treatment proves the role of histone acetylation in the expression regulation of high-molecular-weight glutenin genes. Plant Mol Biol Rep 36:750–763. https://doi.org/10.1007/s11105-018-1117-8

Éva C bZIP transcription factors repress the expression of wheat high molecular weight glutenin subunit genes in vegetative tissues (ms in preparation)

Guo W, Yang H, Liu Y, et al. (2015) The wheat transcription factor TaGAMyb recruits histone acetyltransferase and activates the expression of a high-molecular-weight glutenin subunit gene. Plant J 84:347–359. https://doi.org/10.1111/tpj.13003

Gupta N, Gupta AK, Singh NK, Kumar A (2011) Differential expression of PBF Dof transcription factor in different tissues of three finger millet genotypes differing in seed protein content and color. Plant Mol Biol Rep 29:69–76. https://doi.org/10.1007/s11105-010-0208-y

Holden NS, Tacon CE (2011) Principles and problems of the electrophoretic mobility shift assay. J Pharmacol Toxicol Meth 63:714. https://doi.org/10.1016/j.vascn.2010.03.002

Li X, Gao S, Tang Y, et al. (2015) Genome-wide identification and evolutionary analyses of bZIP transcription factors in wheat and its relatives and expression profiles of anther development related *TabZIP* genes. BMC Genomics 16:976. https://doi.org/10.1186/s12864-015-2196-7

Makai S, Tamás L, Juhász A (2014) Distinct regulatory modules identified in the promoters of wheat *Glu-1* genes suggest different regulatory mechanisms. bioRxiv https://doi.org/10.1101/011635

Makai S, Éva C, Tamás L, Juhász A (2015) Multiple elements controlling the expression of wheat high molecular weight glutenin paralogs. Funct Integr Genomics 15:661–672. https://doi.org/10.1007/s10142-015-0441-4

Mathew IE, Das S, Mahto A, Agarwal P (2016) Three rice NAC transcription factors heteromerize and are associated with seed size. Front Plant Sci 7:1638. https://doi.org/10.3389/fpls.2016.01638

Moreno-Risueno MÁ, González N, Díaz I, et al. (2008) FUSCA3 from barley unveils a common transcriptional regulation of seed-specific genes between cereals and Arabidopsis.

Plant J 53:882–894. https://doi.org/10.1111/j.1365-313X.2007.03382.x
Pharr DM, Stoop JMH, Williamson JD, et al. (1995) The dual role of mannitol as osmoprotectant and photoassimilate in celery. HortScience 30:1182–1188. https://doi.org/10.21273/hortsci.30.6.1182
Ravel C, Martre P, Romeuf I, et al. (2009) Nucleotide polymorphism in the wheat transcriptional activator *Spa* influences its pattern of expression and has pleiotropic effects on grain protein composition, dough viscoelasticity, and grain hardness. Plant Physiol 151:2133–2144. https://doi.org/10.1104/pp.109.146076
Ray DK, Gerber JS, Macdonald GK, West PC (2015) Climate variation explains a third of global crop yield variability. Nat Commun 6:5989. https://doi.org/10.1038/ncomms6989
Shewry PR, Halford NG (2002) Cereal seed storage proteins: Structures, properties and role in grain utilization. J Exp Bot 370:947–958. https://doi.org/10.1093/jexbot/53.370.947
Suzuki A, Wu CY, Washida H, Takaiwa F (1998) Rice MYB protein OSMYB5 specifically binds to the AACA motif conserved among promoters of genes for storage protein glutelin. Plant Cell Physiol 39:555–559. https://doi.org/10.1093/oxfordjournals.pcp.a029404
Tanaka M, Kikuchi A, Kamada H (2008) The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. Plant Physiol 146:149–161. https://doi.org/10.1104/pp.107.111674
Zemlyanskaya EV, Wiebe DS, Omelyanchuk NA, et al. (2016) Meta-analysis of transcriptome data identified TGTCNN motif variants associated with the response to plant hormone auxin in *Arabidopsis thaliana* L. J Bioinform Comput Biol 14:1641009. https://doi.org/10.1142/S0219720016410092